International Journal of Electrical and Electronics Engineering Research (IJEEER)
ISSN(P): 2250-155X; ISSN(E): 2278-943X
Vol. 5, Issue 1, Feb 2015, 87-94
TJPRC Pvt. Ltd.

MOTION ESTIMATION BASED ON NOISE REDUCTION IN COLOR VIDEO USING LMMSE AND ANTISOTROPIC DIFFUSION FILTERING

SARASWATHY C, RAMYA R & VIVEKANANTH

Assistant Professor, Department of ECE, K. S. Rangasamy College of Technology, Tiruchengode, Tamil Nadu, India

ABSTRACT

The denoising method is to remove the additive white Gaussian noise in the color video sequences based on inter frame and inter color correlation prediction is presented. Noises are added uniformly and non uniformly then the temporal correspondence is achieved through the joint-RGB motion estimation to estimate motion between the frames. Pure interframe, pure intercolor and combined interframe-intercolor correlations are predicted and compared between them. Anisotropic diffusion is a technique aiming at reducing image noise without removing significant parts of the image content, typically edges, lines or other details that are important for the interpretation of the image. Linear minimum mean square error filtering (LMMSE) and antisotropic -diffusion filter is applied to the current noisy observation as well as the interframe and intercolor prediction to estimate every noise free color component. Finally the denoised performance are evaluated by the quality metric, color peak signal to noise ratio (cPSNR).

KEYWORDS: Video Denoising, Sub-Pixel Motion Estimation, Motion Compensation, Noise Free Estimator, LMMSE, Antisotropic Filter

INTRODUCTION

Video signals are often contaminated by noise during acquisition, storage and transmission. Removing and reducing noise in video signals is highly desirable, as it can enhance perceived image quality, increase compression effectiveness, facilitate transmission bandwidth reduction, and improve the accuracy of the possible subsequent processes such as feature extraction, object detection, motion tracking and pattern classification. Since the origins of the degradations are numerous and diverse, the overall noise contribution is often modeled as an additive white Gaussian noise independent from the original uncorrupted video sequence. Video denoising algorithms are classified based on whether they are implemented in the pixel domain or wavelet domain. In pixel domain approaches, the filtering is performed directly on pixel intensities[6]. Transform-domain denoising techniques come slightly later than pixel-domain ones and become an active research area[4]. For every nonoverlapping blocks in the current frame, finding the corresponding prediction blocks in the reference frames by motion estimation achieves blocking artifacts and high denoising error [5]. When the color signal is concerned, the strong intercolor correlation is also the essential characteristic so, the interframe and intercolor prediction in overlapping blocks would overcome these problems and achieves low denoising error. The color peak signal to noise ratio evaluates the denoising performance.

VIDEO DENOISING BASED ON INTERFRAME & INTERCOLOR PREDICTION

The video denoising consists of various steps are as follows,

The first step of the process would be preprocessing, which involves conversion of video into frames to denoise

the video in frame by frame manner.

The next step deals with the joint-RGB motion estimation and compensation between the frames in block by block manner.

The next process would involve in applying linear minimum mean square error filtering to the current frame's block and compensated frame's block.

Repeat the above two steps until all the blocks get denoised.

The final step involves aggregating all block-wise estimates to get a final estimate of the whole frame.

FRAMEWORK

Video signal is basically any sequence of time varying images. A still image is a spatial distribution of intensities that remain constant with time, whereas a time varying image has a spatial intensity distribution that varies with time. Video signal is treated as a series of images called frames. An illusion of continuous video is obtained by changing the frames in a faster manner which is generally termed as frame rate. The total number of pixels per second in video can be calculated by the product of the width, height and frames per second. If the larger portion of the sequences are occupied by smoother region means, the frames per second will be more. In the chair video sequence the frames per second is forty three instead of thirty because of the smoother regions in this sequence. In the framework, the frames are separated into three are one current frame and two reference frames. One of this reference frame is denoised previous frame and another reference frame is noisy future frame. To denoise the current noisy frame, the two reference frames are considered. The overall framework of this technique is shown in the Figure 1. The current frame is processed block by block with the fixed block size of $K_x \times K_y$. The distance between two adjacent vertical blocks as well as the distance between two horizontal blocks are D_x and D_y respectively. To divide the current frame into overlapping blocks the distance between two vertical blocks and the distance between two horizontal blocks should be smaller than the block size so that the blocking artifacts can be suppressed.

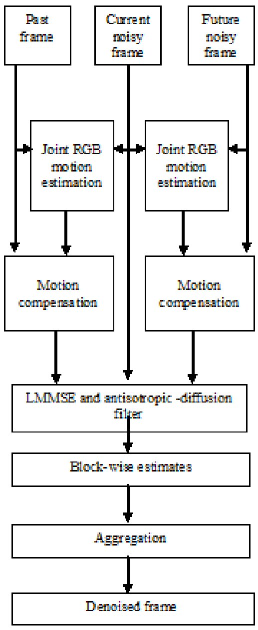


Figure 1: Overall Framework

Then the joint-RGB motion estimation and compensation followed by linear minimum mean square error filtering in block by block manner to estimate every noise free color component. Since all the three color components in the multiple frames as the input to the filtering process, it takes advantages of the interframe and intercolor correlation directly. Block-wise estimates comprises a complete representation of the current frame. All those block-wise estimates are aggregated to get a denoised frame. Block wise processing will be discussed in the following section.

- **Motion Estimation and Compensation**

Considering the three frames, the motion is estimated between the current frame and denoised previous frame as well as between the current frame and the noisy future frame before applying the linear minimum mean square error filtering. The motion estimation is to predict the motion vectors in the RGB space. Block matching algorithm is used to predict the motion vectors[4]. This algorithm is simple and easy to implement and effective over a wide range of video content.

A block matching algorithm is a way of locating matching blocks in a sequence of digital video frames. In this, current frame is divided into equal size blocks are called source blocks. The objective of block matching is to find a candidate block in the search region of the reference frame which suits best to the source block. Based on block distortion measure, the displacement of the best matched block is described as the motion vector to the block in the current frame. The best match can be evaluated by the cost function such as sum of absolute difference (SAD), sum of squared difference (SSD). The lowest SAD and SSD estimates the best position.

$$D = \sum_{i=1}^{M} \sum_{j=1}^{N} (X_{ij} - Y_{ij}) \tag{1}$$

Where D is the difference between two block of pixels, M and N are rows and columns of the block, X_{ij} and Y_{ij} are the current block and the reference block.

$$Z = \text{abs}\,(D) \tag{2}$$

The weighting vectors of the three color components are α_R, α_G and α_B. The sum of squared difference (SSD) is given by,

$$SSD = \sum [\alpha_R(Z^2) + \alpha_G(Z^2) + \alpha_B(Z^2)] \tag{3}$$

The sizes of image block and the search area have a strong impact on the performance of the motion estimation result. A small size block offers a good approximation to the moving object, but it also produces a large amount of redundant motion information data.

When pixels from several frames get blended together there is a ghosting artifacts. To avoid this ghosting artifacts, the estimated motion is compensated by reshaping the predicted motion vectors. Motion compensation describes a picture in terms of the transformation of a reference picture to the current picture.

- **Interframe and Intercolor Correlation**

In many existing video denoising algorithms, only the interframe correlations has been exploited[1]-[4]. When concerning the color signal, this technique also exploits the intercolor correlation for potential denoising gains. The interframe and intercolor correlation has decomposed into three types of correlations as depicted in Figure 2. are pure intercolor, pure interframe and combined interframe-intercolor correlations.

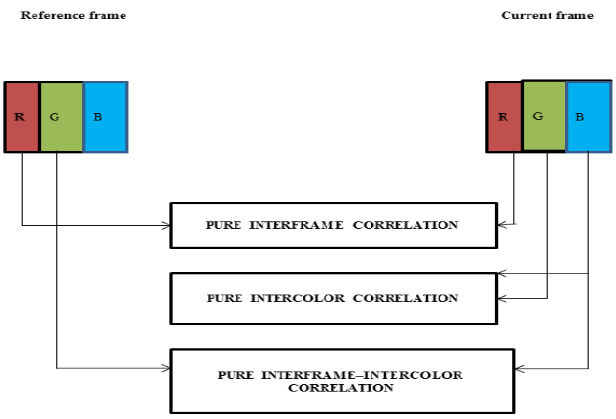


Figure 2: Interframe and Intercolor Correlation

The first one is the pure interframe correlation which is the correlation between the same color component of the current block in the current frame and its motion compensated block in the neighbouring reference frame.

The second one is the pure intercolor correlation which is the correlation between any pair of color components in the same frame.

The third one is the combined interframe-intercolor correlation which is the correlation between one color component of the current block in the current frame and a different color component of its motion compensated block in the reference frame.

The three types of correlation can be analysed by two measures: the correlation coefficient and the mean squared error (MSE). The correlation coefficients and the mean squared error are calculated block by block. If the correlation coefficient is near to +1, then there exists stronger positive correlation between the current frame and the motion compensated frame. It is define by,

$$r = \frac{\sum_m \sum_n (A_{mn} - \bar{A})(B_{mn} - \bar{B})}{\sqrt{\left(\sum_m \sum_n (A_{mn} - \bar{A})^2\right)\left(\sum_m \sum_n (B_{mn} - \bar{B})^2\right)}} \quad (4)$$

where m and n are rows and columns of the frame, r is the correlation coefficient, A_{mn}, B_{mn} are the current frame's block and the motion compensated reference frame's block, $\bar{A}$ and $\bar{B}$ are the mean of a A and B. Smaller the MSE value, the closer the fit is to the data. Using equation (1) and (2) It is given by,

$$MSE = \frac{1}{MN}\sum(Z^2) \quad (5)$$

- **Linear Minimum Mean Square Error Filter**

Linear Minimum mean square error is the algorithm which minimizes mean squared error to denoise the video sequence. This filter is applied in this technique after completion of joint RGB motion estimation. Thus from the current frame and the motion compensated frame, the local block is extracted. Then separate the R, G and B color components from each block. From the separated color component, made the noise observation. and the linear minimum mean square error filtering is applied to estimate every noise free color component and repeat till all the local blocks in the frames get

processed. The flow diagram of this filter is shown in the Figure 3. The formulation of the noisy observation is as follows

$$[V_R, V_G, V_B] = [y_R, y_G, y_B] + [n_R, n_G, n_B] \quad (6)$$

Where V_R, V_G, V_B are the noisy observation of y_R, y_G, y_B and n_R, n_G, n_B are the noise components in $K_x \times K_y$ block. The formulation of the color noise removal by the linear minimum mean square error filtering filtering is as follows

$$Y_R = W_R^T (V_R - U_R) + y_R \quad (7)$$

$$Y_B = W_B^T (V_B - U_B) + y_B \quad (8)$$

$$Y_G = W_G^T (V_G - U_G) + y_G \quad (9)$$

Where U is the mean of V for all the three components the weighting vectors for the filtering of the red, green and blue components are $W_{R,}$ $W_{B,}$ $_{WG}$ and its value are choosed as one for the three components.

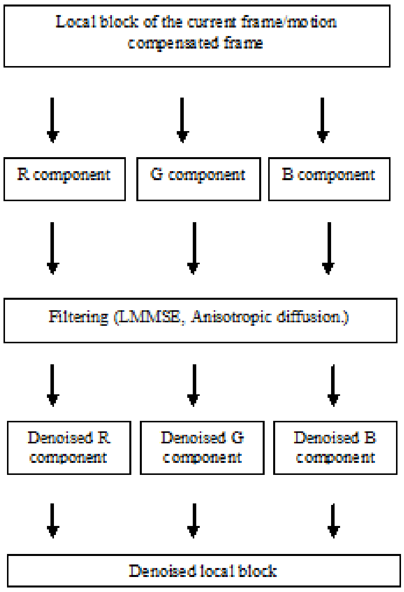


Figure 3: Flow Diagram of the Linear Minimum Mean Square Error Filter

EXPERIMENTAL RESULTS

- **Input Frames**

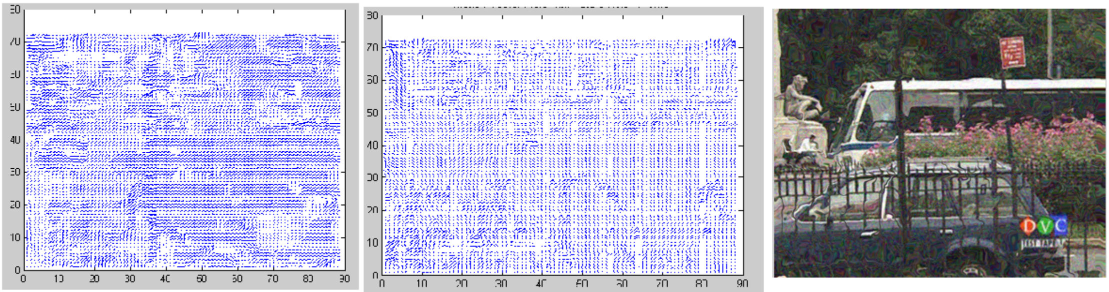
Experiments are conducted on two video sequences named Bus and Chair. The frames per second of the Bus video sequence is thirty whereas, for the chair video sequence it is forty three instead of thirty this is due to larger portion of the sequence occupied by smother regions. Here the past frame is considered as tenth frame of the video sequences, current frame is eleventh frame and future frame is twelth frame. The additive white Gaussian noise is added to the current frame and the future frame uniformly to each color components with the variance 0.02 and nonuniformly with the variances 0.001, 0.003 and 0.008 as shown in the Figure 4.

- **Motion Vector Field and Compensation**

The motion estimation and compensation between the current and past/future frame is shown in the Figure 5. The block

(a) (b) (c)

Figure 4: Input Frames of the Sequence Bus. (a) Past Frame, (b) Current Noisy Frame, (c) Future Noisy Frame



(a) (b) (c)

Figure 5: Motion Estimation and Compensation of the Sequence Bus (a) Motion Vector with Past Frame. (b) Motion Vector with Future Frame (c) Compensated Frame with Past Frame

size is choosed as 8 × 8 with $D_x = D_y = 4$. These choices gives a good trade-off between the motion estimation accuracy and robustness towards noise.

- **Correlation Coefficients and MSE**

To analyse the interframe correlation, intercolor correlation and interframe-intercolor correlation, the correlation coefficient and mean squared error are measured block by block and the values are averaged and reported as shown in the Table 1. Among the three correlations, pure intercolor correlation has the highest correlation coefficient and lowest mean squared value.

Table 1: Interframe & Intercolor Correlation Coefficients

Sequence	Correlation Coefficients		
	Pure Interframe Correlation Coefficient	Pure Intercolor Correlation Coefficient	Interframe Intercolor Coefficient
Bus	0.76	0.85	0.78
Chair	0.73	0.80	0.70

Table 2: Mean Squared Error

Sequence	Mean Squared Error		
	MSE of Pure Interframe Correlation Prediction	MSE of Pure Intercolor Correlation Prediction	MSE of Interframe Intercolor Prediction
Bus	52.8	36.4	53.3
Chair	35.2	32.8	38.3

- **Filtered Denoised Frame**

After the completion of motion estimation and compensation, the linear minimum mean squared error filtering is applied to current frame's block and motion compensated frame's block of pixels The estimation of each and every color component gives a denoised frame as depicted in the Figure 6. Instead of applying linear minimum mean squared error filtering to all the color components as a whole, the filtering to each individual components gives a good performance

- **Color Peak Signal to Noise Ratio**

The denoised performance for uniform noise and nonuniform noise are evaluated by the color peak signal to noise ratio (cPSNR) to each three color components and compared between them with the average color peak signal to noise ratio as shown in the Table 3

In uniform noise, the color peak signal to noise ratio be same for three color components whereas in nonuniform noise, the color peak signal to noise ratio is different for every three color components. Higher the color peak signal to noise ratio when compared to the original clean frame higher the denoise performance

Figure 6: Denoised Frame of the Sequence Bus

Table 3: cPSNR Comparison of the Bus & Chair Sequence

Sequence	Without Noise	With Noise	
	cPSNR	Uniform Noise (cPSNR)	Nonuniform Noise (cPSNR)
Bus	20.2174	25.9792	25.9374
Chair	17.9053	24.4960	24.4201

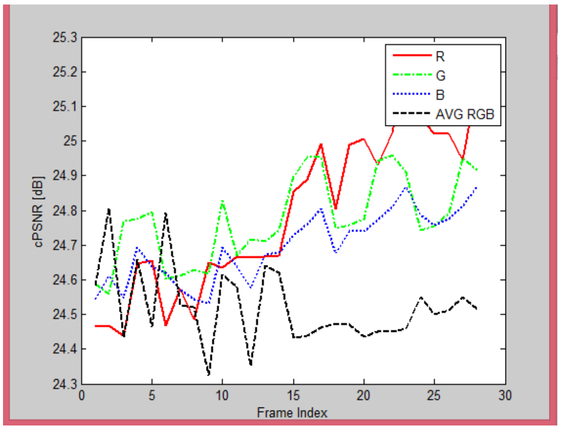


Figure 7: cPSNR plot of the Sequence Chair

CONCLUSIONS

The additive white Gaussian noise can be effectively reduced in video sequences. The joint-RGB motion estimation is applied to estimate motion between the current frame and the reference frame which exploits the interframe and intercolor correlation in the RGB space. Noise reduction is achieved through the LMMSE and antisotropic -diffusion filtering. The interframe and intercolor prediction can be done in overlapping block-wise manner which improves the denoising performance. The three types of correlations can be anlysed by the two measures such as correlation coefficient and Mean Squared Error (MSE). The correlation coefficient value is high for intercolor correlations and MSE value is low for intercolor prediction. The denoising performance can be evaluated by using the quality metric, color Peak Signal to Noise Ratio (cPSNR) and its value is high for denoised frame when compared to original frame. The future work will be done with the transforms to achieve noise reduction and comparison with the above parameters.

REFERENCES

1. Hossein Rabbani, Mansur Vafadust, “Image/video denoising based on a mixture of Laplace distributions with local parameters in multidimensional complex wavelet domain,” Signal Processing 88 (2008) 158–173, July 2007.
2. Gijesh Varghese and Zhou Wang, “Video Denoising Based on a Spatiotemporal Gaussian Scale Mixture Model,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 7, July 2010.
3. Yubing Han, Rushan Chen, “Efficient video denoising based on dynamic nonlocal means,” Image and Vision Computing 30 (2012) 78-85, Dec. 2011.
4. Florian Luisier,Thierry Blu, and Michael Unser, “SURE-LET for Orthonormal Wavelet-Domain Video Denoising,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 20, no. 6, June 2010.
5. J. Dai, O. C. Au, W. Yang, C. Pang, F. Zou, and X. Wen, “Color video denoising based on adaptive color space conversion,” in Proc. IEEE Int Symp. Circuits Syst., June 2010.
6. S.-W. Lee, V. Maik, J. Jang, J. Shin, and J. Paik.Noise-adaptive spatio-temporal filter for real-timenoise removal in low light level images.IEEE Tran.Consumer Electronics, 51(2):648–653, 2005.
7. L. Guo, O. C. Au, M. Ma, and Z. Liang, “Temporal video denoising based on multihypothesis motion compensation,” IEEE Transactions on Circuits and Systems for Video Technology, vol. 17, no. 10, Oct. 2007.